A Comparative Study of Various Approaches for Forecasting Vegetable Prices

Ancy Philip, Elaiya Bharathi, Sainyam Kapoor
Department of Computer Science and Engineering
College of Engineering, Guindy, India
mail:ancyphilip1994@gmail.com

Keywords: ANN, ARIMA, HoltWinters, ETS, Vegetable Prices, Inflation, Linear

ABSTRACT:

This study predicts the wholesale price of 5 vegetables using data sourced from Dindigul agricultural wholesale market, during January 2014 to June 2015. The HW smoothing, ARIMA, ETS models are chosen for the analysis and their prediction power are compared. The results found that ARIMA is better given the linear nature of vegetable prices for a short period of time.

1.INTRODUCTION :

The price of vegetable changes rapidly and the change has a great impact on consumer's daily life. They become the major measure of inflation, change in living and consumption indexes. The precise and correct prediction of values are needed to assist traders,farmers and consumers in the decision of buying, producing the right crop,vegetable. Scientifically collecting the trend of vegetable prices and other agri-oriented commodities prices are of great importance to consumers. We aim to address the issue faced by consumers by predicting prices of important vegetables and help government pay attention to soaring prices. In addition this prediction also aims at helping the farmers grow the right crop at right time based on weather history.

Five important vegetables of Indian meal was considered for our analysis. We localised our data to Dindigul district of Tamil Nadu to eliminate any geographical and weather influences in prediction. In this paper we have analysed different models and given a comparison of the of the prediction power of Autoregressive Integrated Moving Average **(**ARIMA), Holt Winters (HW) and Exponential Smoothing State Space Model (ETS). Six months worth of data was given as input for model to train.

The rest of the paper is organised as follows. Section 2 describes about Time Series Modeling. Section 3 describes about the Data Preparation, Section 4 describes about Implementation and results. Section 5 describes about comparison among the models to find the best model. Section 6 concludes the paper and shows direction for the future work

2. TIME SERIES MODELLING

2.1 ARIMA (Autoregressive Integrated Moving Average) model

ARIMA are in general most general class of model used for forecasting time-series. The model is applied in cases where the data does not tend to be stationary.It can be made stationary by a technique called "differencing". ARIMA models are known to show efficient capability in forecasting at short-terms. It constantly outperformed complex structural models in short-term prediction [3].

2.2 Exponential smoothing state space model

Exponential smoothing is used to smoothen the data so that many window functions act as low-pass filter to remove noise. Exponential smoothing add exponentially decreasing weights as the observations get older. In case of moving averages weights assigned are equal to 1/N in exponential smoothing.[4][5]

2.3 Holt-Winters Smoothing

The Holt-Winter (HW) smoothing is employed if time-series data show increasing or decreasing trends and the slope of the graph varies with the change of time.

3.DATA PREPARATION

3.1 DATA COLLECTION

Data of vegetables prices from Dindigul market has been recorded at http://www.oddanchatramvegetablemarket.com. Time period of 15 Jan 2015 to 15 Sept 2015 was considered. Scraping was done and unique attributes of each vegetable were collected. Attributes like vegetable name, weight and price were considered for analysis. From the data collected, 5 vegetables of prime importance to the typical Tamil Nadu family household were selected based on consumption rate and price - Onion, Okra, Beetroot, Cabbage, Potato.

Table 1 indicates EDA on the data set. Onion has the maximum mean price of Rs. 31 with okra at the lower end with Rs. 19. The large price difference between the minimum, maximum price indicates the importance of the prices of these vegetables to common man. Figure 1 indicates the plot of price of vegetable prices (along y axis) versus time (along x axis). As can be confirmed from the figure, Onion has the highest peak at Rs. 64. This sudden rise can be attributed to the climatic changes.

TABLE 1. EXPLORATORY DATA ANALYSIS

Vegetable	Min	Max	Mean	S.D.
Onion	16	64	31.251	8.355
Beetroot	10	48	28.561	8.985
Cabbage	12	34	20.409	3.897
Okra	8	36	19.571	6.252
Potato	18	46	27.248	7.598

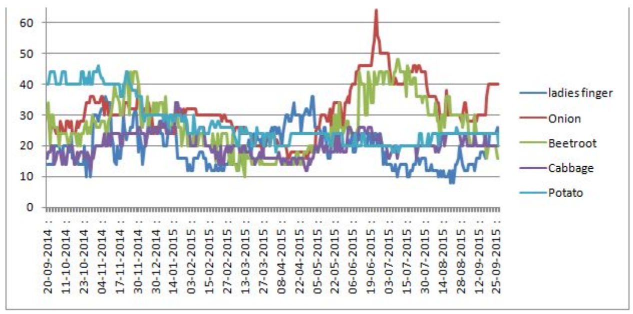


FIGURE 1. Price variation of vegetables

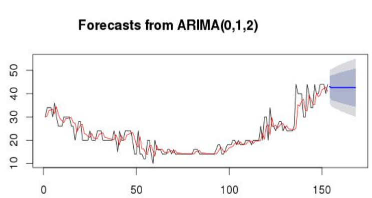
3.2 DATA PARTITION:

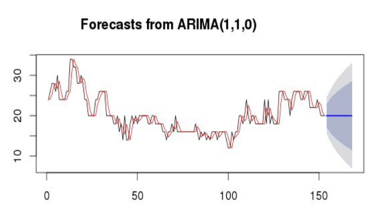
The entire data is partitioned into training data and testing data. The period of **"2015-01-15" to "2015-06-30"** was considered for training data and the period of **"2015-07-01" to "2015-07-15"** was used as testing data. 167 data points of the number series has been chosen for training and the next 15 days has been chosen for testing.

4.IMPLEMENTATION AND RESULTS:

Figures indicating the plot of price (along y axis) vs. time period (along x axis) has been included in each model. Mean Absolute Error (MAE), Mean Percentage Error (MPE), MAPE (Mean Absolute Percent Error), RMSE (Root Mean Square Error) between the predicted and the actual values represent the indicators of prediction power. This information has been represented in tables.

4.1 ARIMA





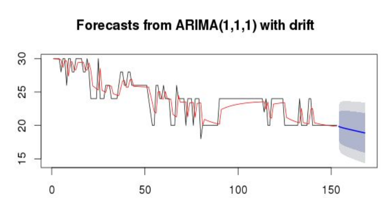
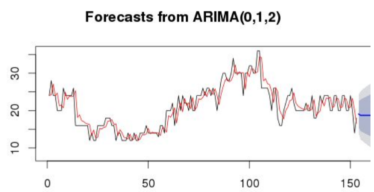


Fig 2.1 ARIMA- Beetroot **Fig 2.2. ARIMA-Cabbage** **Fig 2.3. ARIMA-Potato**



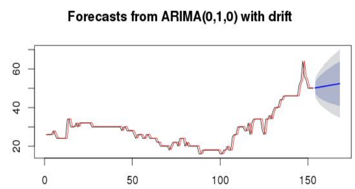


Figure 2: Plot of price prediction for vegetables -ARIMA model

Fig 2.4 ARIMA-Okra

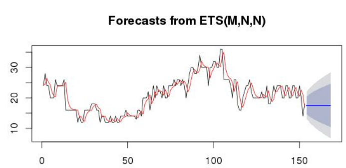
Fig 2.5 ARIMA- Onion

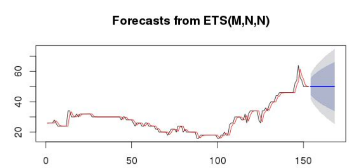
Table 2. Evaluation metrics for Training and testing data in ARIMA model

Train	ME	RMSE	MAE	MPE	MAPE
Beetroot	0.2022	3.5983	2.4514	-1.2969	10.8620
Cabbage	-0.036	2.2592	1.6075	-0.9324	7.9788
Okra	-0.0758	2.6677	2.0323	-1.6449	10.1290
Onion	0.0001	2.3068	1.2736	-0.4607	4.4326
Potato	-0.0129	1.6691	1.1819	-0.4630	5.0506

Test	ME	RMSE	MAE	MPE	MAPE
Beetroot	0.5347	3.1291	2.5755	0.7062	6.0690
Cabbage	-1.2857	2.0701	1.2857	-7.7380	7.7380
Okra	-5.4686	5.6007	5.4686	-42.560	42.560
Onion	-8.8984	9.7673	8.8984	-21.914	21.9144
Potato	0.2037	0.9089	0.8089	0.8506	4.2127

4.2 Exponential Time Series





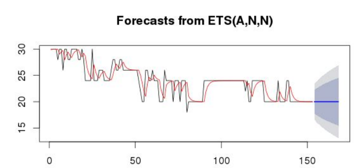
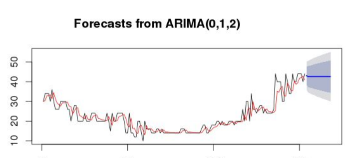


Fig. 3.1- ETS- Okra

Fig. 3.2- ETS- Onion

Fig. 3.3 ETS- Potato



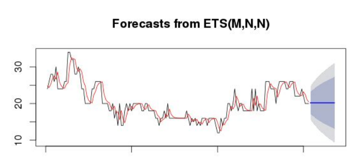


Figure 3 : Plot of price prediction for vegetables -ETS

Fig . 3.4 ETS-Beetroot

Fig. 3.5- ETS-Cabbage

Table 3. Evaluation metrics for Training and testing data in ETS model

Train	ME	RMSE	MAE	MPE	MAPE
Beetroot	0.6195	3.6255	2.5992	0.878	11.4817
Cabbage	-0.0528	2.2760	1.6843	-1.102	8.3483
Okra	-0.0785	2.7370	2.0842	-1.623	10.2947
Onion	0.2067	2.2963	1.3197	0.206	4.5340
Potato	-0.1372	1.7469	1.1312	-0.988	4.8301

Test	ME	RMSE	MAE	MPE	MAPE
Beetroot	-6.5066	7.7724	6.5066	-15.731	15.7319
Cabbage	-1.4797	2.1959	1.4797	-8.783	8.7831
Okra	-4.2693	4.4404	4.2693	-33.454	33.4546
Onion	-7.7286	8.5318	7.7286	-19.061	19.0613
Potato	-0.4289	0.926	0.4289	-2.382	2.3829

4.3 Holt-Winters Smoothing

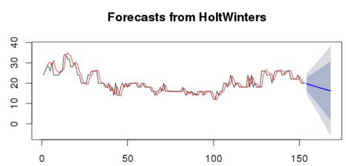


Fig. 4.1 HW-Cabbage

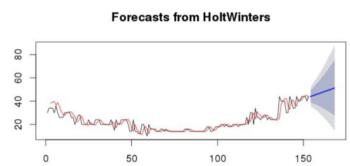


Fig. 4.2 HW- Beetroot

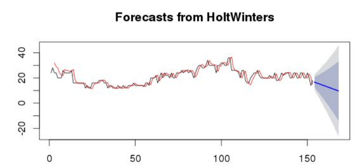


Fig. 4.3 HW-Okra

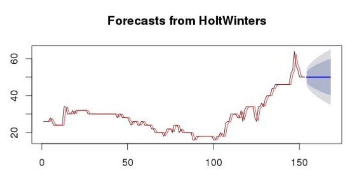


Fig. 4.4 HW-Onion

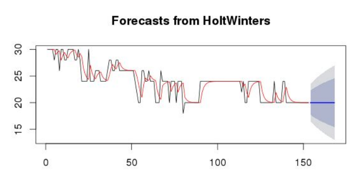


Fig. 4.5 HW- Potato

Figure 4: Plot of price prediction for vegetables -HW

Table 4. Evaluation metrics for Training and testing data in HW model

Train	ME	RMSE	MAE	MPE	MAPE
Beetroot	-0.4281	2.9432	1.8732	-2.253	8.3313
Cabbage	-0.1664	1.9875	1.3768	-1.039	6.9090
Okra	-0.1614	2.4919	1.7785	-1.150	8.5500
Onion	-0.1306	1.9836	1.0264	-0.636	3.6809

Test	ME	RMSE	MAE	MPE	MAPE
Beetroot	-4.1958	5.6362	4.2624	-10.339	10.4803
Cabbage	0.6339	2.3953	2.1034	2.3454	11.3646
Okra	-0.2003	2.4032	2.2064	-2.576	16.9059
Onion	-7.7166	8.5209	7.7166	-19.032	19.0326

Potato	**-0.0079**	**1.1606**	**0.7319**	**-0.305**	**3.0997**

Potato	**-0.4289**	**0.9259**	**0.4289**	**-2.382**	**2.3829**

5. COMPARISON AMONG MODELS:

RMSE for 3 of the vegetables is lesser in ARIMA compared to other models. This can be attributed to data showing evidence of non-stationarity. ARIMA is useful in such cases, where an initial differencing step (corresponding to the "integrated" part of the model) can be applied to reduce the non-stationarity.
MAPE is best for all vegetables in Holt-Winters. HW takes additive and multiplicative seasonality into effect thus accounts for all seasonal trends in the data

Table 5. Comparison of RMSE across models

Train set for Vegetables	RMSE-ARIMA	RMSE-ETS	RMSE-HW
Beetroot	3.1291	7.7724	5.6362
Cabbage	2.0701	2.1959	2.3953
Okra	5.6007	4.4404	2.4032
Onion	9.7673	8.5318	8.5209
Potato	0.908977	0.926	0.9259

6. CONCLUSION AND FUTURE WORK:

Among all agricultural produces, vegetables have the greatest price variability and the most unpredictable supply, as they are grown on bare ground . [2] Vegetable prices are controlled by a variety of factors like local temperature, rainfall , river water quality , climatic/seasonal changes, natural disasters, oceanic salinity/heat content , cyclone heat, cultivation area ,imports/exports ,diesel prices of trucks transporting them, taxes, octroi et cetera. Hence, prediction accuracy can be improved by taking all such factors into account. Moreover, accurate predicted values of weather information coupled with vegetable can aid long term prediction of vegetable prices . This can help government take necessary actions in formulating policies. In addition, common man can plan his diet according to his family income based on Recommended Dietary Allowances(RDA).

7.LIST OF REFERENCES:

[1] Ayodele A. Adebiyi, Aderemi O. Adewumi and Charles K. Ayo, “Stock Price Prediction Using the ARIMA Model “,UKSim-AMSS 16th International Conference on Computer Modelling and Simulation, 2014

[2] Sohyun Kim , Kuk Hyun Nam , Cheolho Song and Youngchan Choe, “The Development and Evaluation of Onion and Cabbage Wholesale Price Forecasting Models”, International Journal of Software Engineering and Its Applications Vol. 9, No. 8 , pp. 37-50, 2015

[3] A. Meyler, G. Kenny and T. Quinn, “Forecasting Irish Inflation using ARIMA Models”, Central Bank of Ireland Research Department, Technical Paper, 3/RT/1998.

[4] Holt, Charles C. "Forecasting seasonals and trends by exponentially weighted moving averages." International journal of forecasting 20.1 (2004): 5-10

[5] Winters, Peter R. "Forecasting sales by exponentially weighted moving averages." Management Science 6.3 (1960): 324-342

[6] Kalekar, Prajakta S. "Time series forecasting using holt-winters exponential smoothing." Kanwal Rekhi School of Information Technology 4329008 (2004): 1-13.